

Sinh viên nộp bài một file duy nhất **STT_HọTen_MaSV.html** vào email: **nh tinh.aw@gmail.com**

Cho dataset điểm **PISA** năm 2015 của các học sinh Việt Nam tham gia kỳ thi này. Anh/chị hãy thực hiện các yêu cầu sau:

- 1) **(1.0đ)** Đọc file dữ liệu trên và hiển thị 5 dòng đầu tiên và 10 dòng ngẫu nhiên của bảng dữ liệu.
- 2) **(0.5đ)** Cho biết kích thước của bảng dữ liệu
- 3) **(1.0đ)** Tính các trị thống kê trung bình (*mean*), độ lệch chuẩn (*standard deviation*) và khoảng dữ liệu (*range*) cho các cột điểm thi **pv1MATH**, **pv1READ** và **pv1SCIE**.
- 4) **(0.5đ)** Liệt kê 10 học sinh có điểm Toán (**pv1MATH**) cao nhất.
- 5) **(1.0đ)** Vẽ đồ thị so sánh mối tương quan giữa Boy và Girl của cột **Gender**.
- 6) **(1.0đ)** Thống kê để hoàn thành bảng số liệu dưới đây, từ đó chọn loại đồ thị phù hợp để trực quan hóa dữ liệu trong **Bảng 1**.

Bảng 1

Vùng (REGION)	Số học sinh	Phần trăm (%)
SOUTH		
CENTRAL		
NORTH		

- 7) **(1.5đ)** Trình bày kết quả vào bảng số liệu sau, và hiển thị kết quả trên màn hình

Bảng 2

Vùng (REGION)	URBAN		RURAL		REMOTE	
	N	%	N	%	N	%
SOUTH						
NORTH						
CENTRAL						

- 8) **(1.5đ)** Thay vì trình bày dữ liệu bằng **Bảng 2**, anh/chị hãy chọn và minh họa bằng biểu đồ hợp lý nhất để tóm tắt dữ liệu cho bảng này.
- 9) **(1.0đ)** Liệt kê các học sinh vùng **SOUTH** có điểm **pv1MATH** >500
- 10) **(1.0đ)** Liệt kê các học sinh **URBAN** của vùng **SOUTH** có điểm **pv1MATH**, **pv1READ**, **pv1SCIE** đều lớn hơn 500. Đồng thời kết quả hiển thị chỉ lấy 5 cột (*Region, Area, Gender, pv1MATH, pv1READ, pv1SCIE*)

----- Hết -----

Sinh viên tuyệt đối không trao đổi bài, các bài giống nhau bị điểm 0 Thực hành